Chapter 17

Approaches to Image Abstraction for Photorealistic Depictions of Virtual 3D Models

Jürgen Döllner and Jan Eric Kyprianidis

Hasso-Plattner-Institut, University of Potsdam, Germany
{doellner, kyprianidis}@hpi.uni-potsdam.de

Abstract

In our contribution, we present approaches of automatic image abstraction, applied to images and image sequences derived as views of virtual 3D city models and landscape models. We first discuss the requirements of cartography-specific visualization based on the virtual globe metaphor as well as the specific characteristics and deficiencies of visualization based on photorealism. We introduce a concept that extends the classical visualization pipeline by cartography-specific functionality, object-space and image-space abstraction, which also represent the two principle ways for implementing cartographic visualization systems. Abstraction provides the prerequisites to visually communicate uncertainty, to simplify and filter detailed elements, and to clearly encode displayed information of complex geospatial information. In addition, it offers many degrees of freedom for artistic and stylistic design of cartographic products. Furthermore, we outline general working principles and implementation of an automatic image-space abstraction technique we developed that creates high-quality, simplified, stylistic illustrations from color images, videos, and 3D renderings.

Keywords: Non-photorealistic rendering, image abstraction, virtual 3D city models

G. Gartner, F. Ortag (eds.), *Cartography in Central and Eastern Europe*,
Lecture Notes in Geoinformation and Cartography, DOI 10.1007/978-3-642-03294-3_17,

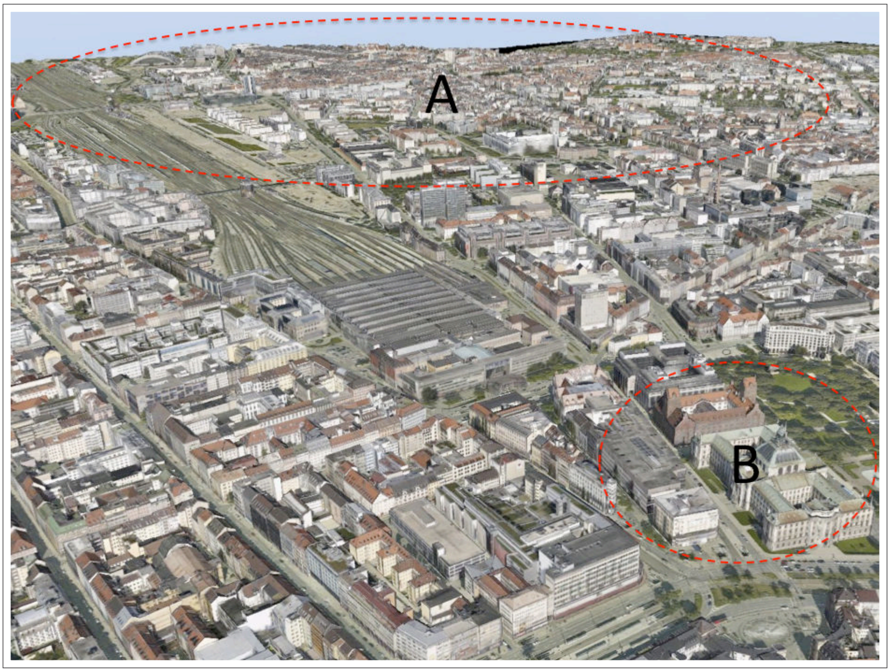

Fig. 17.1. Example depiction with regions of low information value (A) and relatively good visual information value (B)

17.1 Cartographic Visualization Based on the Virtual Globe Metaphor

A growing number of applications and systems use the *Virtual Globe* metaphor as central concept and technology to communicate geospatial information. The base concept is to represent geospatial data as well as georeferenced data by means of virtual 3D models such as virtual globes, virtual city models and virtual landscape models.

These systems already represent key elements in today's information infrastructures and become essential system components in complex workflows in administration and economy (e.g., Google Maps) as well as intuitive, effective user interfaces to geospatial information for the general public (e.g., Microsoft Virtual Earth, Google Earth). As a common characteristic, these systems support naïve geography and thus enhance the specific transmission of geospatial information (Däßler 2002). In addition, communication theories support the assumption that 3D presentation forms establish a naïve information transmission for topographic issues (Egenhofer and Mark 1995).

The ultimate goal, effective communication of geospatial information, has to take into account perceptional, cognitive, and graphical design issues to ensure a clear and efficient understanding of information contents (Jobst 2008). These principles known for 2D depictions also apply to dynamic 3D presentations. In particular, 3D presentations of virtual city models are faced by massive occurrence of occlusion and massive amounts of visual details. Due to that, they demand for specific perceptional, cognitive, and graphic designs compared to well-known, straightforward photorealistic depictions, which often expose large areas of "dead pixels" *(Figure 17.1),* i.e., areas that do not anymore have real information contents. In contrast, illustrative or, more general, non-photorealistic approaches typically achieve better content representation and transmission *(Figure 17.2).*

17.2 City Models and Landscape Models and Their Photorealistic Depiction

Virtual city models and landscape models are based on a multitude of 2D and 3D geospatial information sources such as building and site models, relief models, transportation models, vegetation models, and water models; an overview of principle components has been defined by CityGML, an OGC standard for the exchange of virtual 3D models (Kolbe et al. 2005). The availability of data is not only improved by progress made in remote sensing (e.g., laser scanning and automatic building reconstruction) but also enforced by the increasing compatibility and fusion of data from GIS, CAD/ACE, and BIM (Hagedorn & Döllner 2007). Here, virtual city models and landscape models become general-purpose frameworks for managing, integrating, and using complex, heterogeneous geospatial information.

Image data represents a major part of the data that constitute virtual city models and landscape models – they represent the key elements that enable photorealistic

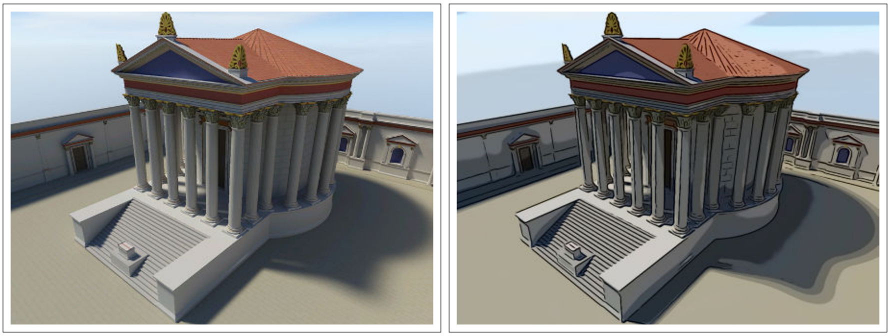

Fig. 17.2. 3D model of a temple of Roman Cologne (Maass et al. 2008): (a) Photorealistic rendering. (b) NPR rendering to communicate missing evidence

depictions. These data sources include, for example, oblique images, orthophotos, and georeferenced photography imagery. In general, in a preprocessing step, the data is filtered, projected, transformed, and assigned to objects of these models, stored frequently in terms of model geometries and textures.

Apart from image data generated by remote sensing, procedurally generated image data represent another important category of photographic data for creating and managing design and appearance of virtual city models and landscape models. The images are generated based on rules that combine pre-defined samples of photographic data (e.g., organized by catalogues for material and surface textures) and synthesize textures for a given object surface (e.g., (Wellmann 2008)).

Photorealistically designed virtual city models and landscape models are generated more and more by automatic processes (Frueh et al. 2004). For example, a collection of oblique images together with a corresponding 3D building model can be automatically processed such that for each visible façade of all building models façade textures are generated. Each façade texture is synthesized by combining parts of the original oblique images according to quality criteria that take into account distance, resolution, occlusion, and error-reducing heuristics about typical façade outfits (Lorenz & Döllner 2006). Insofar we can assume that in the near future the process of creating and updating the geospatial data underlying a virtual city model or landscape model can be automated to a large extend; for base models up to CityGML level-of-detail 2, most likely, this process will be fully automatic.

The photorealistic visualization of virtual city models and landscape models is further enforced by development directions of computer graphics hardware, which concentrate on real-time photorealism. The implementation of systems managing and using massive photorealistic virtual city models and landscape models takes advantage of the texturing capabilities of computer graphics hardware. For example, a number of approaches recently developed allow systems to use an almost unlimited amount of georeferenced texture data within 3D geovirtual environments based on texture atlases (Buchholz 2006).

17.3 Inherent Deficiencies of Photorealistic Visualization

Photorealistic visualization is faced with a number of deficiencies inherent to photorealism. As key characteristic, photorealism wants to represent objects and phenomena by means of the real appearance. In contrast, cartographic visualization relies on abstraction as a fundamental principle to achieve effective communication of complex information. Among the weaknesses of photorealistic visualization we can identify:

- Complex geometry: Models rarely have uniform resolution (e.g., building models derived from cadastral database versus CAD-based models) nor does the resolution correspond in general to the needs and tasks of the user (e.g., detailed CAD models used for an overview of a city).

- Limited texture resolution: The appearance and design of a whole virtual city model suffers from varying and limited resolutions of textures derived by photographic data
- Limited texture quality: The texture contents also expose weaknesses inherent to photorealistic imagery. For example, aerial photography frequently contains non-required objects such as parking cars and pedestrians).
- Fixed lighting and shading: Another inherent weakness of photorealistic imagery represents fixed lighting and shading encoded in photographic imagery. Moreover, they can hardly be removed from the original data.
- Occlusion and perspective distortion: Dependent on the viewpoint, occlusion can strongly impact the visibility of objects. In addition, the perspective projection typically used for 3D viewing leads to distortions and varying scales within a single depiction.

While being attractive in many application areas that benefit from photorealistic visualization (e.g., urban planning, environmental monitoring), photorealism implies strong limitations in the context of cartographic information communication.

17.4 Towards Cartographic Visualization

By *cartographic visualization* we refer to the visualization of geospatial information using concepts and techniques from cartography. Predominantly cartographic visualization relies on abstraction as the most powerful principle to effectively communicate complex geospatial information.

To implement abstraction within the visualization pipeline, two principle approaches can be distinguished: object-space abstraction and image-space abstraction. Both need to be considered independently but have to operate closely at the implementation level. *Figure 17.3* illustrates the corresponding prototypic system architecture.

17.4.1 Object-Space Abstraction

Object-space abstraction refers to an abstraction process that transforms the original 2D and 3D geospatial objects and relations into an intermediate 2D and 3D representation used for rendering. This process typically takes place at the filtering/selection phase of the visualization pipeline.

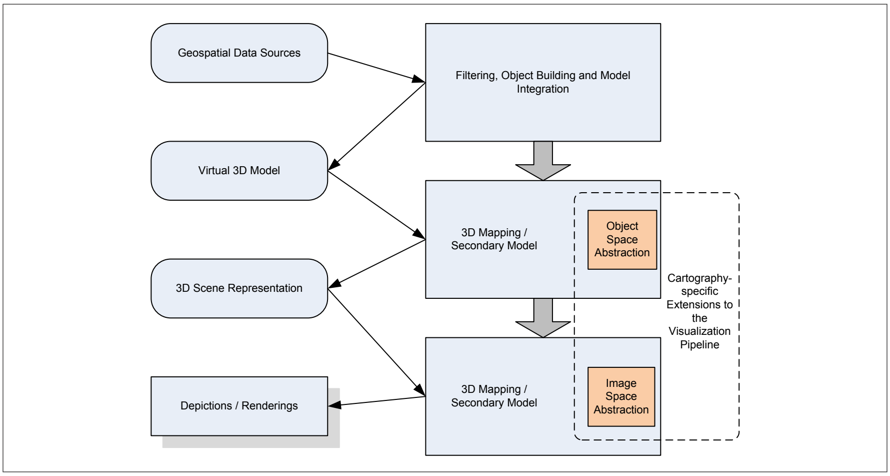

Fig. 17.3. Model of the visualization pipeline extended by cartography-specific components. The data (left column) is transformed through three processing stages (right column) into the final depictions. Object-space abstraction respectively image-space abstraction form part of the corresponding processing stages.

A prominent type of object-space abstraction represents generalization, which in the scope of GI science (Hake et al. 2002) means the process of deriving abstract representations of spatial information subject to a given scale of the communication medium and the intended purpose and tasks to be supported. The process relies on, for example, selecting, combining, reducing, transforming, merging, and enhancing graphical representations; it can also "be understood as a process of representation of the real world by different models" (Cecconi 2003), i.e., derivations of secondary models.

For virtual city models and landscape models, generalization means to derive multi-scale variants. For example, the technique introduced by (Glander & Döllner 2007a) implements the cartographic generalization operators clustering, aggregation, and accentuation. An example of a generalized virtual city model is shown in *Figure 17.4*. It performs the generalization in four steps: 1) City model components are clustered based on the cell structure. 2) For each cell, the weighted average height is calculated, which is also used to automatically identify outliers. 3) Free space is subtracted from the cells such as in the case of outliers or vegetation areas. 4) The modified cells are extruded to building blocks; vegetation areas and outliers are modeled or, respectively, integrated separately. The cell structure can be derived from a given 2D hierarchical network (e.g., hierarchical street and river networks) for which polygonal 2D cells are computed. Each object of the virtual city model is uniquely assigned (and possibly split or re-modeled) to a cluster. The generalized model consists of these representations – that is, abstraction is performed in the scope of the object-space and achieved by generalizing the geospatial data. As key advan-

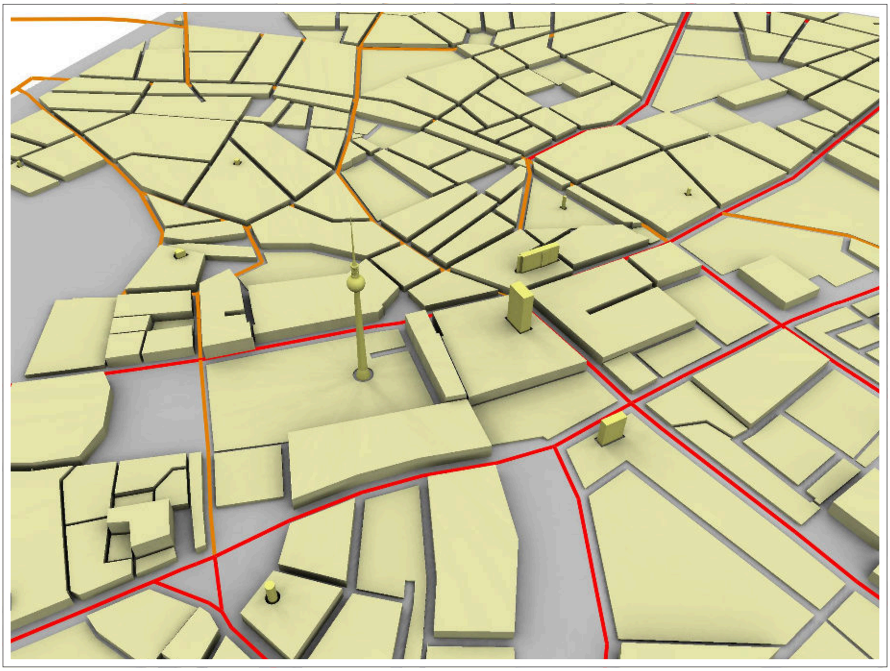

Fig. 17.4. An automatically generalized 3D city model of Berlin, generated by the cell-based method of (Glander & Döllner 2007a)

tage, generalized virtual city models improve the perceptual quality of displayed spatial objects, provide better insights into structure and hierarchy underlying the city model, and facilitate 3D orientation and usability for models with large extend.

Applications and systems based on the Virtual Earth metaphor require object-space abstraction due to manifold reasons such as

- to produce models with homogeneous level-of-detail and, thereby, to integrate heterogeneous geospatial model parts;
- to adapt the level-of-detail of model parts to the current viewpoint;
- to adapt the information density to the needs of users and their tasks.

Object-space abstraction by its nature is non-trivial to implement because it has to automatically analyze, interpret, and synthesize complex 3D geometry, i.e., complete new, consistent secondary models have to be generated. Its automatic implementation in the scope of 3D geospatial data is still in its infancy. Even in the scope of 2D geospatial data the implementation of object-space abstraction techniques represents a major challenge for automation. In particular, techniques must be able to compute results in near real-time to be used in interactive applications and need to define not only discrete levels-of-details but also the transformation in between those levels.

17.4.2 Image-Space Abstraction

Image-space abstraction refers to abstraction applied to the depiction of geospatial representations. It forms part of the rendering process. In contrast to object-space abstraction, it can be implemented by graphics processing, performed in real-time, and applied independently from the underlying geo-objects' representations. Insofar, it is complementary to object-space abstraction and concentrates on the visual design used to render the contents of virtual city models and landscape models.

With respect to cartographic visualization, image-space abstraction demands for dedicated rendering techniques that implement specialized coloring, shading, illumination, and projection techniques. As a simple example of image-space abstraction, consider edge enhancement (Nienhaus & Döllner 2003). It can be implemented by post-processing generated images, applying convolution filtering that enhance color contrasts in the given image. Edge enhancement represents an effective way to improve the perception of the shape of a depicted object.

An example of a more sophisticated image-space abstraction technique is described in *Section 17.5.* It applies complex image analysis and filtering and can be used directly to transform common photorealistic images of virtual city models and landscape models into illustration-like depictions.

17.4.3 Dependencies Between Abstraction Approaches

A cartographic visualization system has to carefully select and co-ordinate object-space and image-space techniques. Object-space abstraction techniques in general will be independently applicable from the 3D rendering system – they operate within the modeling space and, therefore, do not make any assumptions about the rendering technology. Their results may be useful also for analytic computations as well.

There are also hybrid abstraction approaches that require a coordinated implementation. As an example consider an approach that enhances 3D landmarks in perspective views (Glander et al. 2007b). Landmarks represent elements of outstanding importance for user orientation because they facilitate navigation and exploration within large virtual city models and landscape models. The image-space/object-space abstraction technique for enhancing landmarks emphasizes the landmark objects by improving their visibility with respect to their surrounding areas and the current 3D viewing settings; emphasizing is achieved by scaling the landmarks' geometry according to an importance function while simultaneously squeezing their corresponding neighborhood regions *(Figure 17.5).* To reduce visual artifacts caused by this multi-scale presentation, e.g., geometry intersections, the surrounding objects of each landmark are adapted according to a deformation field that encodes the

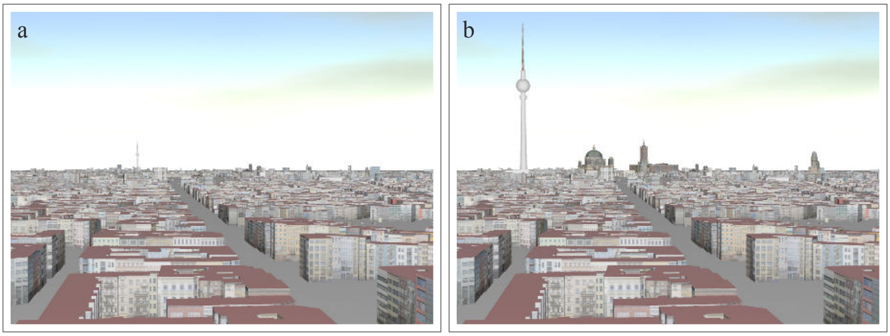

Fig. 17.5. Enhancement of multiple landmarks in a virtual 3D city model (Glander et al. 2007b). Comparison between standard (a) and enhanced rendering (b)

displacement and scaling transformations. An individual weight coefficient can be defined that denotes the landmark's importance. To render a collection of weighted landmarks within a virtual city model or landscape model, the technique accumulates their associated, weighted deformation fields in a view-dependent way. With landmark emphasizing the technique improves the perceptual and cognitive quality of cartographic depictions of 3D models. Coloring and shading of landmarks are emphasized by the image-space part of the technique's implementation.

In an ideal system, both object-space and image-space abstraction techniques would be implemented separately and interchangeably. Still there is a need for concepts and implementations that show how such an interoperable schema could work. OGC as well as all major GIS vendors still do not define dedicated, open standards for these system components – a topic that is at least addressed to a certain degree by the ongoing OGC working group on portrayal services.

17.5 Automatic Image Abstractions for Photorealistic Depictions

We have developed an automatic rendering technique that creates high-quality, simplified, stylistic illustrations from color images, videos, and 3D renderings (Kyprianidis & Döllner 2008). This technique can be added as a general-purpose image-space abstraction technique as post-processing to real-time 3D rendering pipelines and, therefore, is suitable for a variety of interactive 3D applications and systems. It has only minor requirements regarding the contents of virtual city models and their object-space abstractions. For this reasons, it can be used in a straightforward way as a basic building block for cartographic visualization systems and achieves a broad range of illustration styles due to a number of configuration

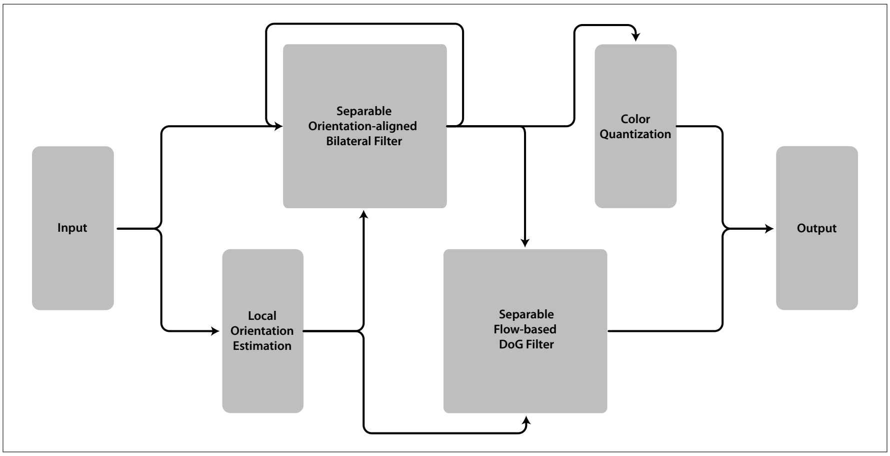

Fig. 17.6. A schematic overview of our image abstraction framework

parameters. Details of the implementation of our technique with focus on recent GPU technology can be found in (Kyprianidis & Döllner 2009).

Our method extends the approach of (Winnemöller et al. 2006) to use iterated bilateral filtering for abstraction and difference-of-Gaussians (DoG) for edge extraction by adapting it to the local orientation of the input. We developed enhancements to these techniques to improve the quality of the output by adapting them to the local orientation of the input. A schematic overview of our framework is shown in *Figure 17.6.* Input is typically an image, a frame of a video, or the output of a 3D rendering. We start with the estimation of the local orientation. Then, the input is iteratively abstracted by using the orientation-aligned bilateral filter. We perform a total of n_a iterations. To the result we apply color quantization. After $n_e < n_a$ iterations, we extract edges from the intermediate result using the separable flow-based DoG filter. In our examples we use $n_e = 1$ and $n_a = 4$. Finally, the extracted edges are superimposed on the output of the color quantization. *Figure 17.7* shows the output of different stages of the algorithm.

17.5.1 Local Orientation Estimation

To represent local orientation we construct a smooth tensor field. From the eigenvectors of this tensor field we derive a vector field that has similar characteristics as the edge tangent flow (ETF) of (Kang et al. 2007), but its computation is much less expensive. Besides gradient calculation, only smoothing with a box or Gaussian filter is necessary. In contrast to that, ETF construction requires several iterations of a nonlinear filter with large filter kernel.

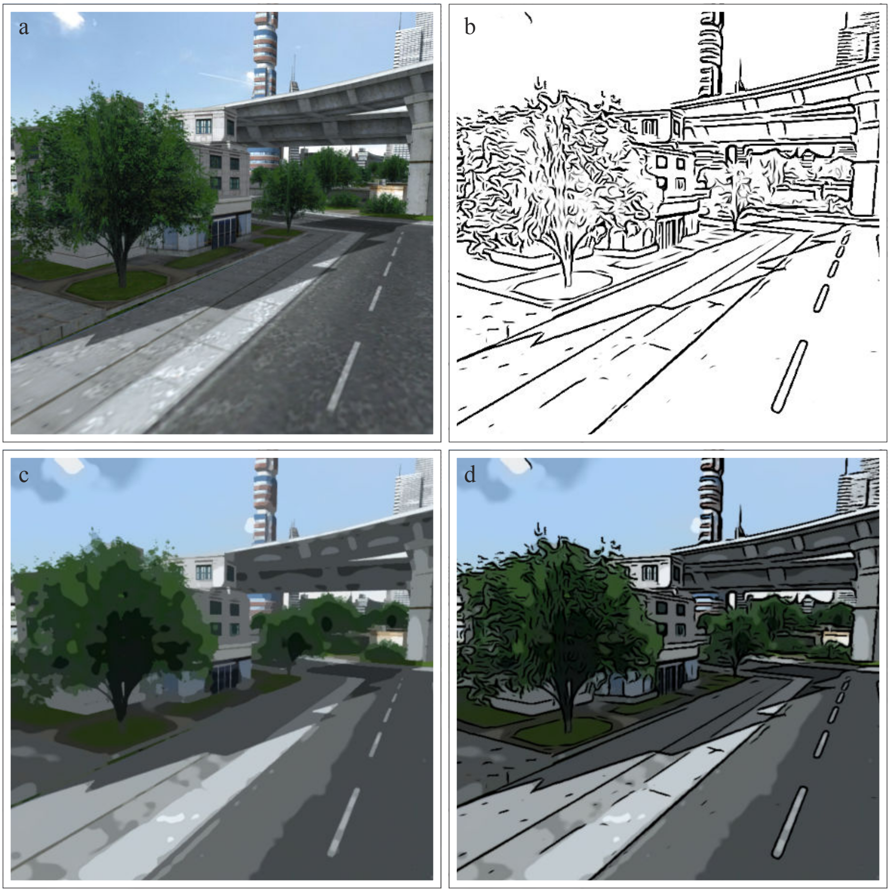

Fig. 17.7. Illustration of immediate results of the algorithm: (a) Original photorealistic rendering. (b) Output of edge detection. (c) Output of bilateral filter with color quantization applied. (4) Abstracted output.

17.5.2 Orientation-Aligned Bilateral Filter

The bilateral filter is a non-linear filter that smoothes images while preserving edges (Paris et al. 2007). It is based on two weighting functions. The first one gives more weight to pixels that are close to the filter center and the second one gives more weight to pixels whose colors are similar to the color at the filter's center. This has the effect that regions of similar color are smoothed, while regions with detail are preserved. For an image *f*, the bilateral filter is defined by:

$$\frac{\sum_{x \in \Omega(x_0)} f(x)\ G_{\sigma_d}\left(\left\|x - x_0\right\|\right) G_{\sigma_r}\left(\left\| f(x) - f(x_0) \right\|\right)}{\sum_{x \in \Omega(x_0)} G_{\sigma_d}\left(\left\|x - x_0\right\|\right) G_{\sigma_r}\left(\left\|f(x) - f(x_0)\right\|\right)}$$

Here, x_0 denotes the center of the filter neighborhood. For the closeness function and the similarity function, we use one-dimensional Gaussian functions:

$$G_\sigma(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{t^2}{2\sigma^2} \right)$$

To achieve a nonlinear diffusion effect, the bilateral filter must be applied recursively. We are therefore interested in a fast computation of the bilateral filter. Computing the bilateral filter is expensive because it is not separable. A simple implementation requires evaluation of the full kernel. Several acceleration schemes have been presented in the research literature. The *xy*-separable bilateral filter (Pham & van Vliet 2005) used by (Winnemöller et al. 2006) suffers from horizontal and vertical artifacts. These artifacts appear in particular when the filter is applied iteratively. Our approach works by first filtering in direction of the gradient and then filtering the intermediate result in perpendicular direction. When applied iteratively our approach does not suffer from horizontal or vertical artifacts and creates smooth output at curved boundaries.

17.5.3 Separable Flow-Based Difference-of-Gaussians

Edges are extracted from the luminance channel after n_e iterations of the bilateral filter. In our examples we typically use a single iteration. This preprocessing is required to avoid the detection of edges that are due to noise. The (Marr & Hildreth 1980) edge detector works by computing the Laplacian-of-Gaussian and detecting the zero crossings in the result. The Laplacian-of-Gaussian can be approximated as the difference of two Gaussians. This variant is called Difference-of-Gaussians (DoG).

DoG edges often look frayed and don't reassemble straight line and curve segments very well. To work around this limitation, (Kang et al. 2007) recently introduced the concept of flow-based difference-of-Gaussians which, compared to DoG edges, create more coherent lines. They replaced the DoG filter by a flow-guided anisotropic kernel whose shape is defined by the ETF. Comparable high-quality results can be achieved by a separated implementation with corresponding reduced computational complexity. We first apply a one-dimensional difference-of-Gaussian filter in direction of the gradient and then apply smoothing along the vector field that we derive from the smoothed structure tensor.

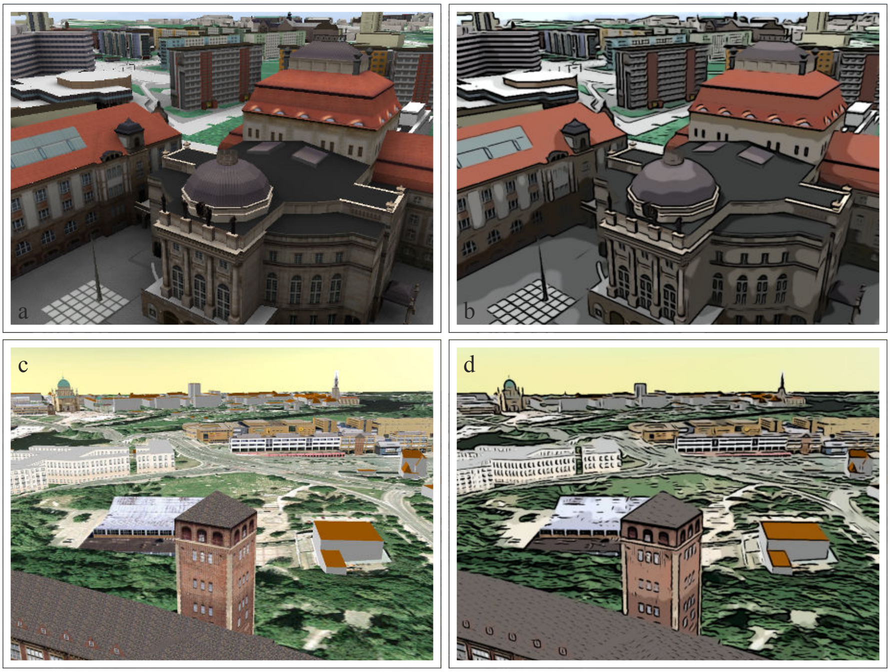

Fig. 17.8. Abstracted depictions produced by our method: (a),(c) Original Images. (b),(d) Abstracted results.

An example set of abstracted depictions produced by our method is illustrated in *Figure 17.8*. The underlying virtual city model is a typical representative of today's frequently found virtual city models with photorealistic design.

17.6 Conclusions

We have discussed essential elements of cartographic visualization, image-space and object-space abstraction as well as a general-purpose technique for image-space abstraction. The presented image-space abstraction helps to achieve a more efficient cartographic display of virtual arbitrary virtual city models and landscape models. Image-space abstraction provides the prerequisites to visually communicate uncertainty, to simplify and filter detailed elements, and to clearly encode the model nature of the displayed information – independently from the degree of generalization of the original underlying models. The technique also shows pragmatic qualities as it can be efficiently implemented on top of today's GPUs and as an add-on to existing rendering pipelines.

In our future work, we would like to investigate object-space abstraction techniques that directly cooperate with the presented image-space abstraction technique. In particular, we would like to separately treat different objects categories (e.g., building models versus vegetation models) and to apply different rendering techniques, respectively.

Acknowledgments

This work was supported by the German Research Foundation (DFG), grant DO 697/5-1.

References

Buchholz H (2006) Real-Time Visualization of 3D City Models. PhD Thesis. Universität Potsdam

Cecconi A (2003) Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping. PhD Thesis. University of Zürich

Däßler R (2002) Visuelle Kommunikation mit Karten. In: Arthur Engelbert, Manja Herlt (ed) Updates – visuelle Medienkompetenz. Königshausen & Neumann

Döllner J, Hagedorn B (2007) Integrating Urban GIS, CAD, and BIM Data By Service-Based Virtual 3D City-Models. 26th Urban Data Management Symposium

Egenhofer M, Mark D (1995). Naive Geography. COSIT 95, Semmering, Austria, In Frank A, Kuhn W (ed) Lecture Notes in Computer Science, Vol 988, Springer, pp 1–15

Frueh C, Sammon R, Zakhor A (2004) Automated Texture Mapping of 3D City Models with Oblique Aerial Imagery. Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission, 396–403, IEEE Computer Society

Glander T, Döllner J (2007a) Techniques for Generalizing Building Geometry of Complex Virtual 3D City Models. 2nd International Workshop on 3D Geo-Information: Requirements, Acquisition, Modelling, Analysis, Visualisation

Glander T, Trapp M, Döllner J (2007b) A Concept of Effective Landmark Depiction in Geovirtual 3D Environments by View-Dependent Deformation. 4th International Symposium on LBS and Telecartography

Gröger G, Kolbe TH, Czerwinski A (2007) City Geography Markup Language (CityGML). OGC Best Practices Document, Version 0.4.0, OGC Doc. No. 07-062. Open Geospatial Consortium

Hake G, Grünreich D, Meng L (2002) Kartographie. 8 Ed Walter de Gruyter, Berlin, New York

Jobst M (2008) Ein semiotisches Modell für die kartografische Kommunikation mit 3D. PhD Thesis. Technische Universität Wien. http://media.obvsg.at/AC05037648

Kang H, Lee S, Chui CK (2007) Coherent Line Drawing. In Proc ACM NPAR, pp 43–50

Kolbe TH, Gröger G, Plümer L (2005) CityGML – Interoperable Access to 3D City Models. Proceedings of the First International Symposium on Geo-Information for Disaster Management. Delft, Netherlands

Kyprianidis JE, Döllner J (2008) Image Abstraction by Structure Adaptive Filtering. In Proc. EG UK Theory and Practice of Computer Graphics, pp 51–58

Kyprianidis JE, Döllner J (2009) Real-time Image Abstraction by Directed Filtering. In ShaderX7 – Advanced Rendering Techniques, Charles River Media

Lorenz H, Döllner J (2006) Towards Automating the Generation of Facade Textures of Virtual City Models. ISPRS Commission II, WG II/5 Workshop, Vienna, www.hpi3d.de/publications

Maass S, Trapp M, Kyprianidis JE, Döllner J, Eichhorn M, Pokorski R, Bäuerlein J, v. Hesberg H (2008) Techniques For The Interactive Exploration Of High-Detail 3D Building Reconstruction Using The Example Of Roman Cologne. Proceedings of 14th International Conference on Virtual Systems and Multimedia (VSMM 2008), pp 223–229

Marr D, Hildreth RC (1980) Theory of edge detection. In Proc. Royal Society London 207:187–217

Nienhaus M, Döllner J (2003) Edge-Enhancement – An Algorithm for Real-Time Non-Photorealistic Rendering. International Winter School of Computer Graphics, Journal of WSCG, 11(2):346–353

Paris S, Kornprobst P, Tumblin J, Durand F (2007) A gentle introduction to bilateral filtering and its applications. In ACM SIGGRAPH courses

Pham TQ, van Vliet LJ (2005) Separable bilateral filtering for fast video preprocessing. In IEEE International Conference on Multimedia and Expo

Wellmann M (2008) Automatische Generierung graphischer Oberflächendetails für Bodenflächen in virtuellen 3D-Stadtmodellen. Master Thesis. Universität Potsdam

Winnemöller H, Olsen SC, Gooch B (2006) Real-time video abstraction. ACM Transactions on Graphics 25(3):1221–1226